REFERENCES

1. A. Axenopoulos ., "I-SEARCH: A Unified Framework for Multimodal Search and Retrieval", in Proc. Future Internet Assembly, 2012, pp.130-141.
2. A. Jaimes ., "Multimedia Information Retrieval: What is it, and why isn‟ t anyone using it?" ACM MIR05, Singapore 2005.
3. A. Hauptmann and M. Christel, "Successful approaches in the TREC video retrieval evaluations", Proc. of ACM Multimedia, Oct. 2004, 668-675.
4. A. K. Jain, and A. Ross, "Multibiometric systems", Communications of the ACM, Special Issue on Multimodal Interfaces 2004, Vol. 47, No. 1, pp.34–40.
5. A. Ross, and A.K. Jain, "Multimodal biometrics: an overview", Proceedings of EUSIPCO, 2004, pp.1221– 1224.
6. A. B. Benitez, J. R. Smith, and S. F. Chang, "MediaNet: a multimedia information network for knowledge representation," in Proc. the SPIE 2000 Conf. Internet Multimedia Management Systems, vol. 4210, 2000.
7. A. Motro, "FLEX: A tolerant and cooperative user interface to databases," IEEE Trans. Knowledge and Data Engineering, 2(2): 231-246, 1990.
8. E. H. Y. Lim ., "Knowledge Seeker - Ontology Modelling for Information Search and Management", Springer Berlin Heidelberg, Vol. 8, pp. 27-36, 2011.
9. E. Kasutani, "Image retrieval apparatus and image retrieving method", US Patent application 2007.
10. Perronnin ., "Large-scale image retrieval with compressed fisher vectors". In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2010.
11. F. Alhwarin ., "Improved SIFT-Features Matching for Object Recognition"; BCS International Academic Conference 2008 – Visions of Computer Science, pp. 179-190; 2008.
12. G. Csurka ., "Visual categorization with bags of keypoints". In Proc. of the ECCV Workshop on Statistical Learning for Computer Vision 2004.
13. G. Hubert and J. Mothe, "An adaptable search engine for multimodal information retrieval". J. Am. Soc. Inf. Sci., 60: 1625–1634. doi: 10.1002/asi.21091, 2009.

14. J. C. Haartsenet. al., "Adaptive Display for Enhancing Audio Playback.", U.S. Patent Application 12/209,300, filed September 12, 2008.
15. J. Kludas, E. Bruno, and S.M. Maillet, “Information fusion in multimedia information Retrieval”, Proceedings of 5th International Workshop on Adaptive Multimedia Retrieval: Retrieval, User, and Semantics, Paris, France, 2007, pp.147–159.
16. J. Foote, “An overview of audio information retrieval,” ACM Multimedia Systems, 7: 2-10, 1999.
17. J. M. Kleinberg, “Authoritative sources in a hyperlinked environment,” in Proc. ACM-SIAM Symposium on Discrete Algorithms, pp. 668-677, 1998.
18. L. Nigay and J. Coutaz, “A generic platform for addressing the multimodal challenge,” in Proc. Int. Conf. ComputerHuman Interaction, pp. 98-105, 1995.
19. M. Davis. S. King. N. Good, and R. Sarvas, “From context to content: leveraging context to infer media metadata”, Proc. of ACM Multimedia. Oct. 2004, 188- 195.
20. M. Flickner, H. Sawhney, W. Niblack, and J. Ashley, “Query by image and video content: The QBIC system,” IEEE Computer, pp. 23-32, 1995.
21. M.M. Kokar, J.A. Tomasik, and J. Weyman, “Formalizing classes of information fusion systems”, Information Fusion 2004, Vol. 5, No. 4, pp.189–202.
22. N. Aggarwal, Dr. Nupur Prakash and Dr. SanjeevSofat, “Mining Techniques for Integrated Multimedia Repositories: A Review”, BVICAM‟S International Journal of Information Technology, Issue 1, JanuaryJuly, 2009 Vol.1 No.1.
23. P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, “GroupLens: an open architecture for collaborative filtering of netnews,” in Proc .of ACM Conf. on Computer-Supported Cooperative Work, pp. 175-186, 1994.
24. R. S. Dubey, R. Choubey and J. Bhattacharjee, “Multi Feature Content Based Image Retrieval”, (IJCSE) International Journal on Computer Science and Engineering, Vol. 02, No. 06, 2010, 2145-2149.
25. R.A. Bolt, “Put that there: voice and gesture at the graphic interface”, Computer Graphics 1980, Vol. 14, No. 3, pp.262–270.
26. R. Sarukkai, “Video search: opportunities and challenges”, Proc. of the ACM workshop on Multimedia information Retrieval, 2005, 3-8.

27. R. Bentley, T. Rodden, P. Sawyer, and I. Sommerville, “An architecture for tailoring cooperative multi-user displays,” in Proc. Conf. Computer-Supported Cooperative Work, pp. 187-194, 1992.
28. R. Elmasri and B. Navathe, Fundamentals of database systems, 2 Edition. The Benjamin/Cummings Publishing Company, Inc., Redwood City, CA, 1994.
29. R. Lempel and A. Soffer, “PicASHOW: pictorial authority search by hyperlinks on the Web,” in Proc. 10 th Int. World Wide Web Conf., pp. 438-448, 2001.
30. S. F. Chang, W. Chen, H. J. Meng, H. Sundaram, and D. Zhong, “VideoQ: an automated content based video search system using visual cues,” in Proc. ACM Multimedia Conf., pp. 313-324, 1997.
31. S. Sarin and W. Kameyama, “Joint Equal Contribution of Global and Local Features for Image Annotation”, CLEF working notes, 2009.
32. S-F. Chang, “Multimedia access and the state of the art and future directions”, Proc. of ACM Multimedia, Nov. 1999, 443-445.
33. S. Brin and L. Page, “The anatomy of a large-scale hypertextual Web search engine,” in Proc. 7 th Int. World Wide Web Conf., pp. 107-117, 1998.
34. S. Oviatt, A. DeAngeli, and K. Kuhn, “Integration and synchronization of input modes during multimodal humancomputer interaction,” in Proc. Int. Conf. Computer-Human Interaction, vol. 1, pp. 415--422, 1997
35. T. Otsuka, A. Utsumi and J. Ohya, “Advanced man-machine interfaces based on computer vision technologies – Recognizing facial expressions and hand gestures,” in IEEE Int. Workshop on Robot and Human Communication, Vol.1, pp.56-63, 1998.
36. Urban, Jana, J. M. Jose, and C. J. V. Rijsbergen. "An adaptive technique for content-based image retrieval", Multimedia Tools and Applications 31 no. 1, 2006, 1-28.
37. G. Hubert and J. Mothe, “An adaptable search engine for multimodal information retrieval”. J. Am. Soc. Inf. Sci., 60: 1625–1634. doi: 10.1002/asi.21091, 2014.
38. Perronnin ., “Large-scale image retrieval with compressed fisher vectors”. In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2015

39. J. Foote, “An overview of audio information retrieval,” ACM Multimedia Systems, 7: 2-10, 2017
40. , . J. C. Haartsen (2018) In this survey, we focus on multimodal document retrieval combining the text and image modalities. Multimodal IR is based on single-modality IR systems, including text retrieval, CBIR, text, Proceedings of EUSIPCO, 2018, pp.1221– 1224.